Does PDS Need to Support a Medium Term "Archive"?

Robert Deen

Paulo Penteado

Fred Calef

Jet Propulsion Laboratory

California Institute of Technology

Motivation

- **PDS is a long-term archive (50+ years)**
  - Holdings are thus limited to static data
- **Flight missions provide lots of services while active**
  - Servers, databases, interactive mechanisms
  - Could be considered "short-term archive"
    - Available during the mission
    - 6-12 months after mission end to support closeout
- **There is nothing formal in between**
- **Thus the question:**

Is there a need for a medium term (5-10 year) archive?

What Kinds of Services?

- **Most modern missions have servers and software to access their data**
  - These would often still be of benefit after the mission ends
- **Some examples will help clarify**
  - These are systems with which the authors are most familiar
  - Many other examples

Marsviewer

- **System of data servers with Java, JavaScript, or iOS client, used for browsing data from MSL, MER, PHX, InSight, Mars 2020. Visualizes image data and derived data – XYZ, range, slope, etc.**
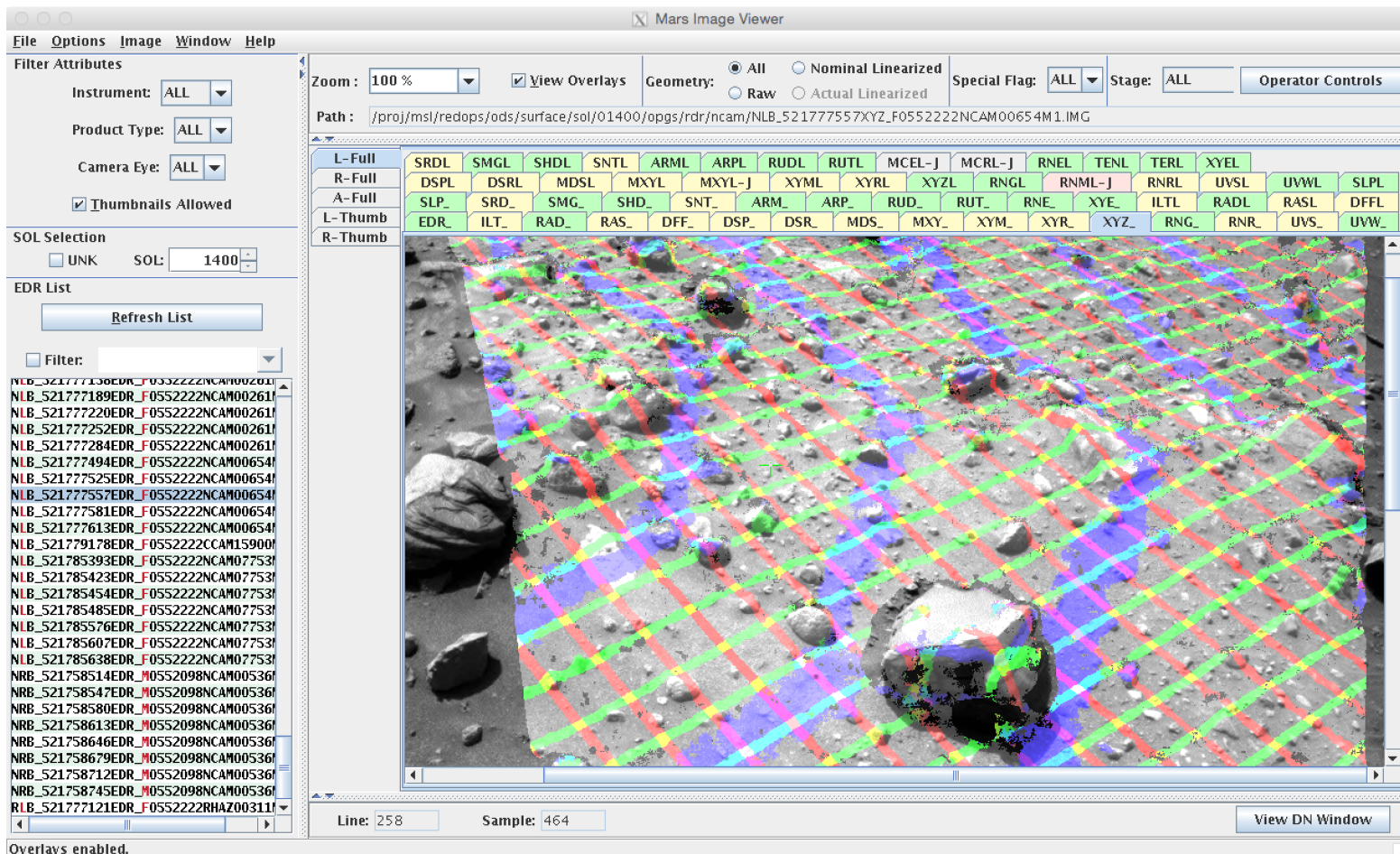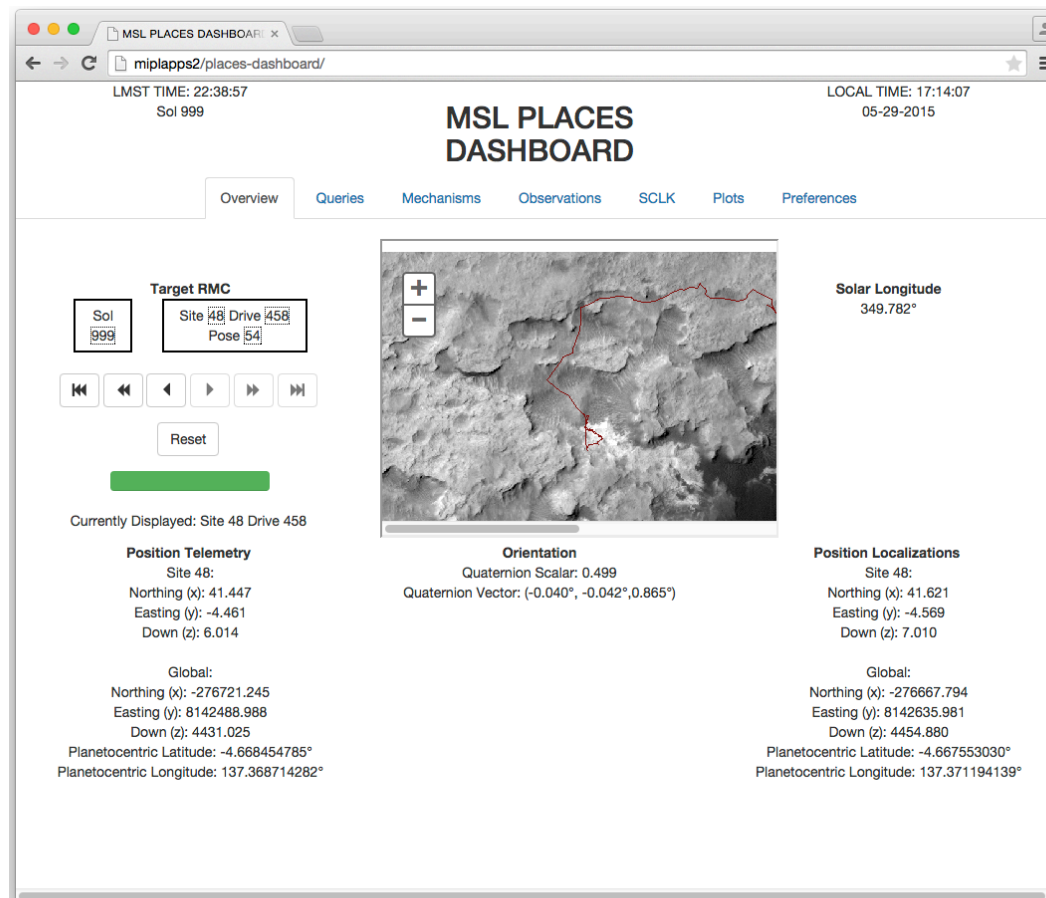
PLACES

- **Active database containing MSL rover localizations (where the rover is). Data is extracted to PDS archive now, but server provides additional functionality: coordinate translation and maintenance of user-created localizations.**
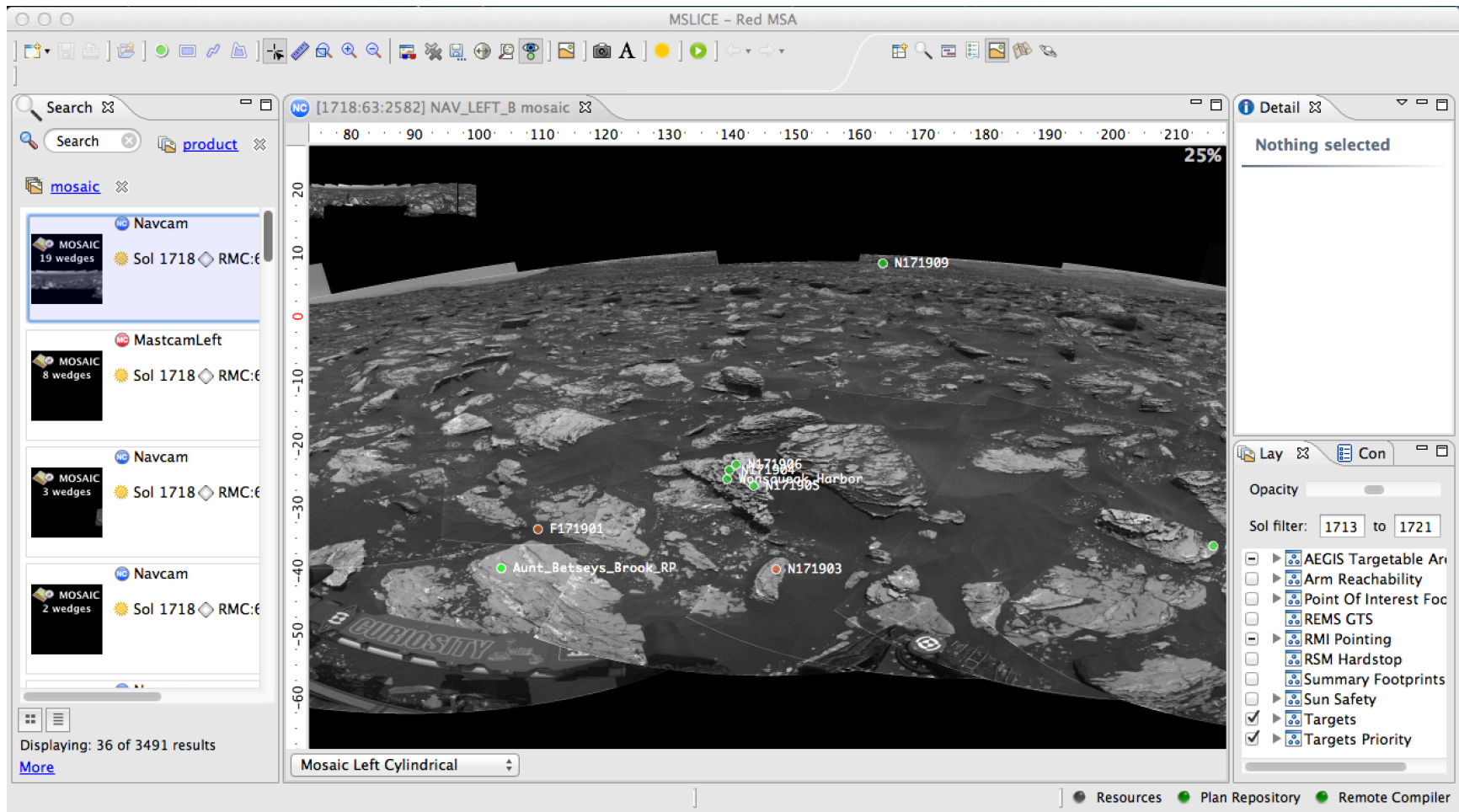
MSLICE

- **Tool used by science team to command instruments, but has a significant visualization component which would be useful after mission end.**
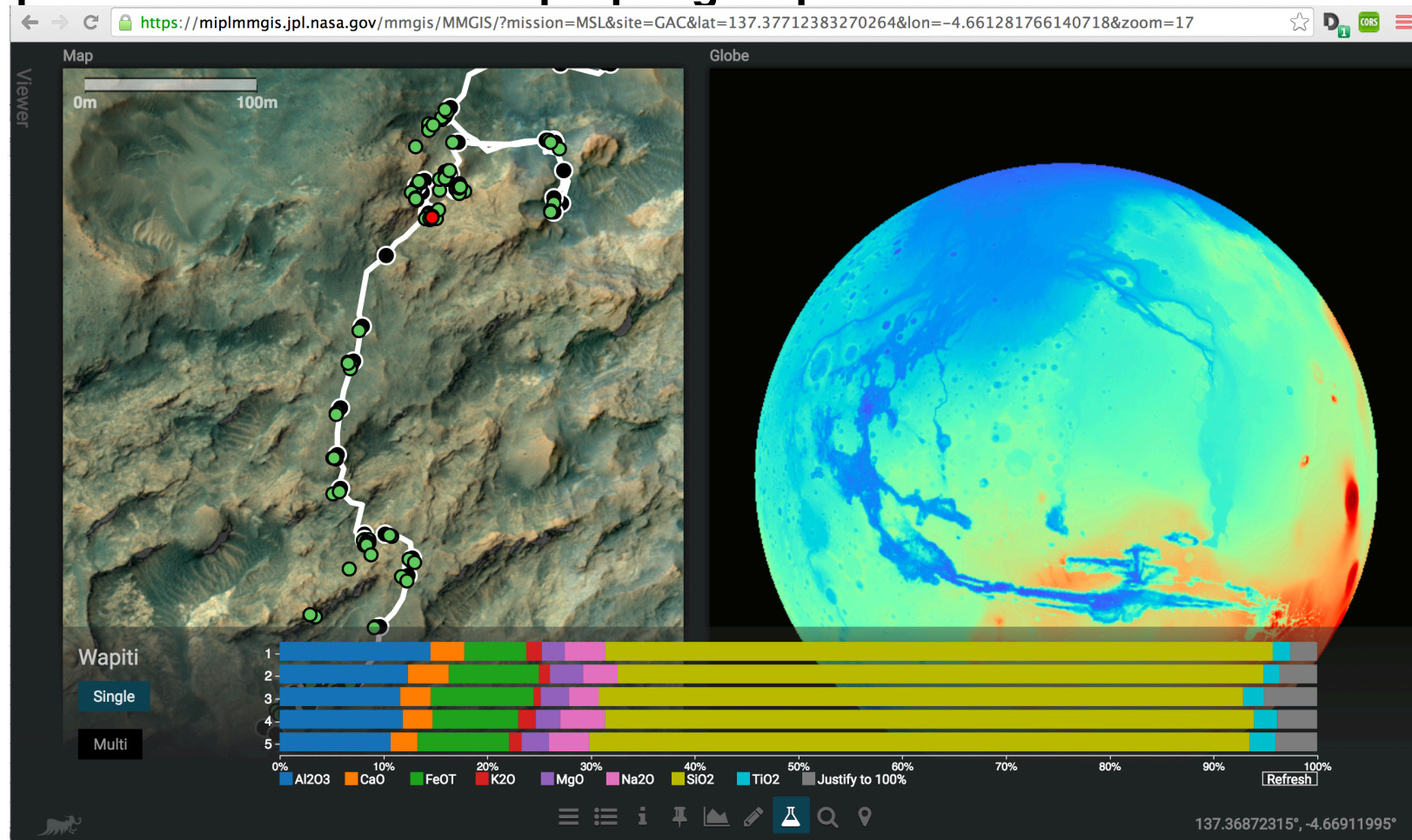
MMGIS

- **GIS system of programs, workflows, and Web interfaces to unite mission basemaps (orbital imagery) with science products and quicklook results in their proper geospatial context.**
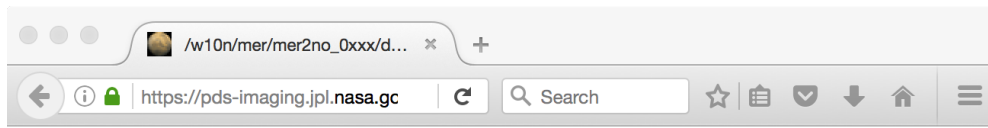
W10N (Webification)
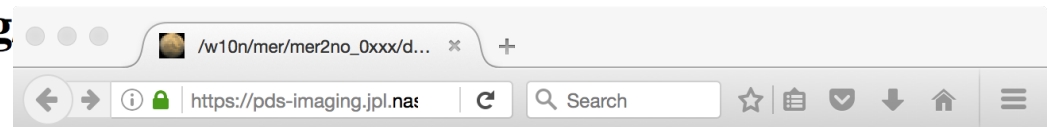
- **General service that provides Web-enabled access to resources. Used as a foundation for the Marsviewer services, among other uses.**

Jmars

- **GIS system deployed by ASU. May have achieved enough success to be self-sustaining, but if ASU pulls the plug, PDS could pick it up.**
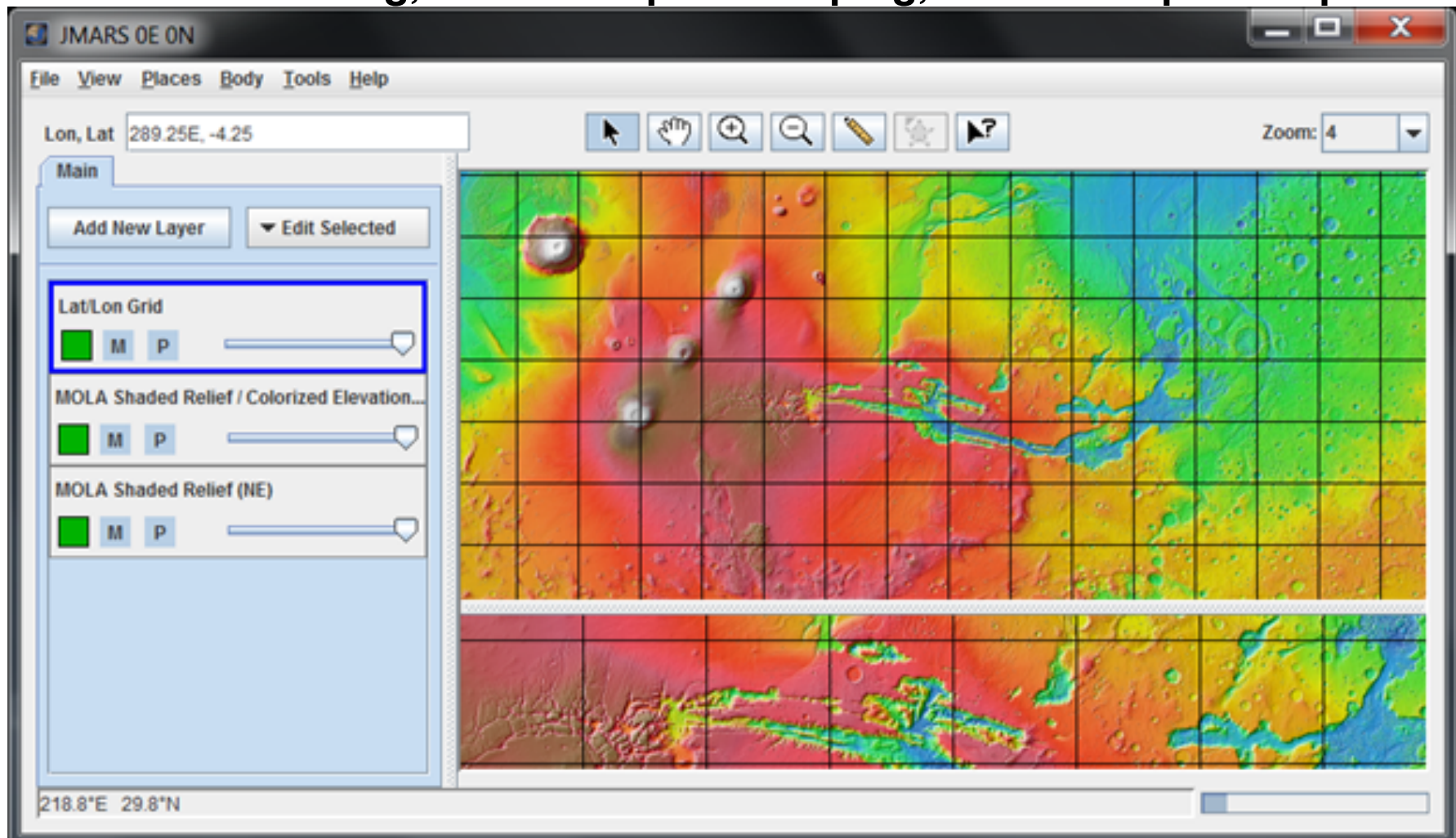
VICAR

- **Image processing software suite (not server). It is 50 years old and continuously maintained, but if support were to cease, PDS could keep offering it for as long as it keeps running.**

Cassini VIMS Data Server

- **Primary source for VIMS science data. Provides calibrated observations on the fly, controlled by user parameters. Maintained by VIMS ops team, but support is likely to end with the mission.**

Titanbrowse

- **Database and visualization service developed for Cassini VIMS, provides queries based on full hyperspectral dataset and arbitrary processing. Potentially applicable to any hyperspectral imaging dataset, not just VIMS.**
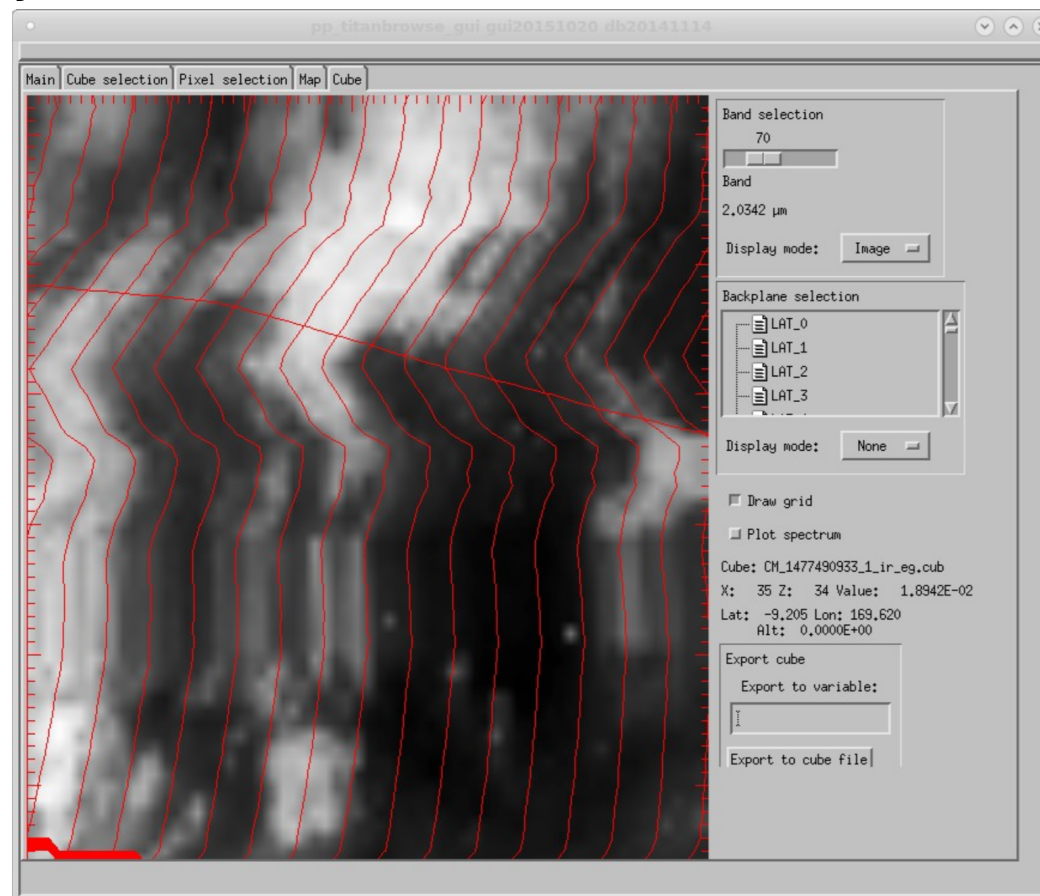
Bit Rot

- **Software tends to "rot" over time**
- **Hardware becomes obsolete**
  - Dedicated frame buffers from the 1980's
  - Array co-processors
  - 9-track tape drives
- **CPU's change**
  - 6502, PDP-11, Z-80, VAX
- **Compilers change**
  - Language choices evolve (Pascal? Ada? COBOL?)
  - Languages evolve
    - Some "K&R" C from the 80's doesn't work with modern compilers
    - Fortran IV vs. 77 vs. 2008
- **Networks evolve**
  - Protocols change over time (security!!)
  - Anyone remember DECnet? BITnet?
- **Operating systems change**
  - Backward compatibility maintained only so long
  - MS-DOS anyone? CP/M? VMS? OS/360?
  - Compatibility issues even between versions of the same OS

The Problem with Software

- **Bit Rot means software is unlikely to work in the future**
  - Unless constantly maintained and updated
- **Simply not credible that today's software will work in 50+ years**
- **What about Virtual Machines?**
  - VM's preserve a given operating environment
  - Could help make old environments available for longer
  - But will VM file formats still be readable in 50 years?
- **50 years is a pipe dream… but 5-10 years is feasible**
  - Especially with VM's

Why Medium Term?

- **Interest in mission data is generally highest during and soon after the mission**

  – Assertion without proof: majority of science results come within 5-10 years after a mission ends

  – Data is likely "mined out" after that

    - Exceptions: change detection, re-visits to same location with new spacecraft

- **This is exactly the timeframe that software can be expected to run with minimal maintenance**

Long-Term is Vital

- **The 50+ year archive of PDS is critical**
  - We are not proposing diluting or de-emphasizing that in *ANY* way
- **Long-term users should be no worse off than they are now**
  - This *must not* be a crutch to skimp on the long-term archive
- **Medium-term simply provides more options**
  - Easier to work with the data
- **Long-term users may benefit as well**
  - Examination of server source code, even if not running, can shed light
  - Future mission visiting the same body may wish to resurrect expired services
    - Preserving them, even in non-running form, is crucial

How Might This Work?

- **We do not claim to have all the answers; we are posing questions.**
- **But in broad brush:**
  - Mission close-out activities include packaging relevant servers and software into VM's
  - PDS nodes deploy these servers
  - PDS nodes keep them running as long as possible without investing too many resources
    - PDS nodes cannot maintain them forever
      - Prohibitively expensive
    - Minor fixes might be possible
  - Original authors, or other interested parties, could update if they choose
    - Open Source contributions
  - No new data, so do not need ingest subsystems of servers
    - Simplifies maintenance
  - At some point, when it breaks, node simply turns it off
    - VM, source code, other artifacts, still available
- **Simply keeping the power on is very helpful**
  - Hard to do once mission ends and nobody is paying the bill

Precedents

- **Imaging Node has already deployed the Java Marsviewer client**
  - And the W10N server it needs
  - Support for Web and iOS versions is planned
  - Marsviewer is being actively maintained for future missions
- **VICAR image processing system**
  - Continuous development for 50 years
  - When a mission ends, code is simply left in place
    - "Benign neglect"
    - It mostly just builds and sits there in case it's needed
    - If it stops building, and the fix is non-trivial, code is obsoleted
      - But generally only when it fails to build
  - Application programs are simpler than servers
    - But concept has proven to be useful

Summary

- The authors feel the answer is:

    YES!!

- PDS does need to support a medium-term archive of software servers and services, which it simply runs as long as it can without requiring major effort to maintain.

- So… how do we make this happen?

- Contact info:
    - Bob.Deen@jpl.nasa.gov
    - Paulo.Penteado@jpl.nasa.gov
    - Fred.Calef@jpl.nasa.gov